

WiseEye Elderly Monitoring System Using Machine Learning

¹Amila Sampath, ²Akila Kavinda, ³Sachintha Gunaratne, ⁴Thushal Shaminda

^{1,2,3,4}Department of Computer Science and Software Engineering, Sri Lanka Institute of Information Technology, Sri Lanka

Abstract - The Wise Eye Elderly Monitoring System, a comprehensive solution for the monitoring and care of elderly people in Sri Lanka, was developed and put into use as part of this study. This system includes monitoring devices and mobile application to reduce the burden on caregivers. The system uses machine learning techniques to enable automatic system activation and deactivation through image classification, assuring user-friendliness and individualized monitoring for each senior. The device also has an automatic accident detection module that uses machine learning techniques to find accidents in real time and identify them, alerting carers who are away from home right immediately by mobile app. Additionally, the system incorporates voice and sound recognition algorithms to improve communication and enable voice-based commands. This functionality will enhance the system's ability to seamlessly identify and classify an incident of emergency based on the voice inputs received. At the very least, the Wise Eye system uses behavioral pattern analysis and machine learning to forecast how active an older person will be, providing details on their daily routines and probable behavioral aberrations. A dataset of real-world scenarios is used to evaluate the system, showing how well it can recognize the elderly, detect accidents, hear spoken commands, and gauge activity levels. The suggested Wise Eye Older Monitoring System offers a clever and thorough solution that increases the older population's safety and well-being while simultaneously lightening the load on caregivers and raising the standard of care delivered.

Keywords: machine learning, identification, detection, elderly monitoring system, senior, device, accident detection, automatic, caregivers, elder, real-time.

I. INTRODUCTION

Elderly care represents a crucial societal concern, especially given the growing global aging population. Key challenges faced by caretakers involve constant vigilance, health monitoring, and safety assurance for the elderly. The conventional care giving methods often impose heavy constraints on both caretakers and elderly individuals. This paper introduces a proper solution for these challenges by

leveraging the capabilities of the Internet of Things (IoT), machine learning, and Automatic Speech Recognition (ASR). This research project aims to design an IoT-based monitoring system integrated with a mobile application, empowered by machine learning algorithms. This innovative system aims to ensure real-time, accurate monitoring of elders, thus enhancing their safety and quality of life. The system employs image classification to activate and deactivate automatically upon identifying the elder's presence, promoting efficiency and ease-of-use. Moreover, the system features an advanced automatic accident detection (AAD) mechanism capable of identifying the precise accident location, triggering immediate alerts to family members or caregivers. This could potentially lead to timely interventions, minimizing the risk of severe injuries. Another component of the system is the ASR system, specifically fine-tuned for elderly voices. This allows the system to understand and respond to voice commands, enhancing the interactive and user-friendly nature of the design. Finally, by conducting behavioral pattern analysis, the system can predict the activity levels of the elderly, providing invaluable insights into their health and well-being. Overall, this study attempts to revolutionize elderly care, proposing a technologically advanced, yet user-friendly solution for the challenges faced in the existing framework.

II. LITERATURE REVIEW

Numerous studies have been conducted on this area. The following section provides an overview of some of the most notable and relevant research findings. These selected studies offer a comprehensive perspective on the existing body of literature.

The challenges related to working with Internet of Things (IoT) technology for healthcare applications, notably remote geriatric monitoring, were discussed in this article [1] by the authors. The limited battery life of sensors caused by continuous data streaming and the dependency on cloud-based computations, which results in latency in real-time monitoring applications, were two significant problems that they acknowledge the IoT faces. The authors created a real-time, secure, and energy-efficient platform to address these issues. Compressive Sensing (CS), a power-saving method, was integrated into the suggested platform to lessen the

computational load on the cloud and reduce data transmission delays. With the use of CS, sensors may recover sensed data with fewer measurements than would otherwise be necessary, prolonging their battery life and lowering data transmission delay. The authors of the study hope to increase the effectiveness and dependability of remote elderly monitoring systems by putting this platform into place, which would lead to better healthcare outcomes and lower costs.

Voice command systems have been recognized as a bridge for those who struggle with modern technology interfaces. While there are many ASR systems for languages like English, languages like Sinhala, which are less resourced, have fewer ASR systems available. The few available systems are not integrated with elderly monitoring systems, forcing Sinhala speakers to use English. Traditional methods, heavily reliant on large datasets, prove difficult for such low-resourced languages. However, transfer learning [2], where a model trained in one language is fine-tuned for another, has shown promise. One study demonstrated high accuracy using this technique with Sinhala and Tamil, even with limited speech data [3].

Approach	Benchmark		Current			
	SVM	6L FFN	TL + SVM	TL + FFN	TL + 1D CNN	TL + 2D CNN
Features	MFCC		DS Intermediate			
Accuracy Sinhala	48.79%	63.23%	70.04%	74.67%	93.16%	92.09%
Accuracy Tamil	29.25%	26.98%	23.77%	35.50%	37.57%	76.30%

Table 1: Summary of results with different approaches and overall accuracy values of Transfer Learning based ASR

The unique characteristics of elderly voices, such as hoarseness and slower speech, necessitate specialized training for ASR models to be accurate. One research effort sought to increase the age range of their dataset by including voices of the "super-elderly," leading to a reduced word error rate. For elderly monitoring systems, where the user base is limited, adapting the ASR system to individual voices is suggested as an effective way to increase accuracy, especially given the limited datasets for elderly Sinhala speakers. Various techniques, including speaker normalization and maximum likelihood linear regression, offer pathways for such adaptation.

In recent years, there has been a lot of interest in the study of estimating activity levels by examining the behavioral patterns of elderly monitoring systems using machine learning. Several studies have investigated using machine learning algorithms to analyze elderly people's behavioral patterns and forecast their activity levels. According to F. Delmastro's article [4], behavioral patterns can be used to keep track of one's health. The article discusses the potential advantages and difficulties of this strategy after reviewing several studies on the subject. Several studies that used

behavioral patterns to track health status are reviewed in this article. These studies have looked at a variety of behaviors, such as social interactions, sleep patterns, and physical activity. The article focuses on the advantages of this method, including the potential to personalize health interventions based on unique behavioral patterns and the capacity to identify changes in health status before symptoms manifest.

III. METHODOLOGY

The "Wise Eye Elderly Monitoring System" is a cutting-edge assistive technology system designed specifically for the elderly. It consists of two main parts: a mobile application and an IoT device equipped with a camera and sensors. This approach was specifically created to improve the supervision and care of senior individuals. By giving caregivers up-to-the-minute information on an elderly person's condition and surroundings, technology helps them respond quickly to their requirements.

The Wise Eye Elderly Monitoring System utilizes existing image and audio datasets to enhance its functionalities. These datasets comprise real-world images capturing daily activities and audio recordings encompassing voice commands and ambient sounds. These datasets enable the system to detect falls, predict activities, respond to voice commands, and identify emergencies. Stringent data privacy measures are in place to ensure confidentiality. Figure 1 shows the system overview diagram.

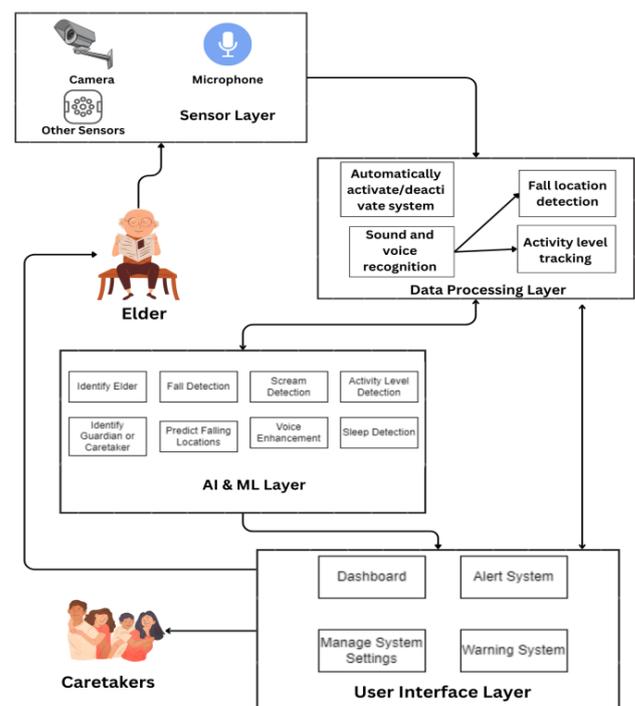


Figure 1: System overview diagram

A) Elder Alone Detection and Device Activation– EAD

A diverse dataset of images was collected, representing various scenarios with and without an elder person alone. After that the collected dataset underwent preprocessing. The images were resized to a consistent resolution, and pixel values were normalized to ensure uniformity. Additionally, the dataset was divided into training and validation sets, allowing for proper evaluation of the model's performance. MobileNet was chosen as the image classification model for this task. The MobileNet model was then fine-tuned using the annotated dataset, adjusting hyperparameters like learning rate, batch size, and the number of training epochs to optimize performance. The performance of the trained model was evaluated using various evaluation metrics, including accuracy, precision, recall, and F1 score. These metrics provided insights into the model's ability to correctly classify images and detect whether an elder person was alone. The results were carefully analyzed to identify any limitations or areas for improvement.

To further refine the EAD model, a process of fine-tuning and iteration was employed. Hyperparameters including batch size and learning rates were adjusted, the dataset was augmented. This iterative approach aimed to enhance the model's performance and address any deficiencies identified during evaluation. Extensive testing and validation were conducted to assess the effectiveness and reliability of the deployed system. Real-world scenarios were simulated, and the system's performance was evaluated based on its accuracy and efficiency in detecting when an elder person was alone. These tests provided valuable insights and performance metrics for further analysis.

B) Automatic Accident Detection of Elderly Person– AAD

The implementation was done by using the 2 action classes (fall and non-fall). The primary and secondary implementations were done with 80% for the train set and 20% for the test set. The implementation was carried out using Inception V3 algorithm under convolution neural network (CNN) [1] by selecting the best approach to the AAD model. The model architecture was defined using a Sequential model from Keras. TensorFlow's data loading utilities were used to load the image dataset [5]. The model was compiled using the Adam optimizer and binary cross-entropy loss function. Training was performed using the fit method of the model with training and validation. During the training phase, the model was trained using the labeled training dataset [6]. The training process involves feeding the images into the model and adjusting the model's parameters to minimize the difference between the predicted fall/non-fall labels and the true labels. The main configurable parameters of the model

architecture were the number of epochs to 150, and the batch size was 128. After the model was trained, it was evaluated using the testing dataset to assess its performance on test data. The accuracy, precision, recall, and other performance metrics calculated to evaluate the model's effectiveness in fall and non-fall recognition.

Future more the system involved motion sensors capturing the movements of an elder. Once motion was detected, the fall detection model would capture a sequence of frames and transmit them to the classification model through the custom-made APIs. The purpose was to identify the occurrence of an accident. If the classification model detected an accident, an emergency alert was promptly sent to the caregiver's mobile application throughout the server in real time. This provided them with the precise location of the elderly's fall. This systematic process ensured a swift response and support in critical situations, allowing caregivers to promptly address the needs of the elderly.

C) Speech Recognition with Enhanced Elderly Voice– EEVR

The initial stage of the research methodology involved the collection of voice data samples. Those are gathered from two distinct groups, namely elderly and non-elderly individuals [7]. A dataset provided by 'Open Speech and Language Resources' was used as the non-elderly data samples. It includes around 185,000 data samples but for this research only 11,500 data samples were used because it takes much time to train the model. This step was crucial to create a broad and diverse dataset, which would be the foundation for training the ASR system. After collecting the voice data samples, the next phase of the research involved the analysis of those audio samples. To facilitate this analysis, Python package called Librosa was used. This tool, familiar to a microscope for sound, allowed checking each audio file in detail. The analytical capabilities of Librosa helped to extract features from the collected audio data.

The research hit a bottleneck when it came to the number of voice data samples from elderly people [8]. Given the scarcity of this data, the methodology had to be adopted. This is where the technique of transfer learning [2] were integrated. This approach resembles how skills acquired in one area can be applied to a different but related area. An ASR model pre-trained on a variety of voices was utilized, and then fine-tuned on a limited set of elderly voices. This technique proved to be an efficient way of enhancing the performance of the EEVR model with reduced data requirements [9]. Once the ASR system had been adequately trained, it was important to ensure accurate understanding of the intent behind the spoken words. To accomplish this, a classification method known as Support

Vector Machines (SVM) was utilized. SVM is known for its efficiency and excellence in identifying patterns in limited data, enabling the system to effectively discern and categorize intents, thus enhancing the system's interaction capabilities.

D) Activity Level Prediction of Elder– ALP

The dataset includes a wide range of actions that are unique to elderly people, such as daily routines, mobility tasks, housework, leisure activities, and self-care procedures. Each image is meticulously assigned to the appropriate action class, ensuring accurate and trustworthy ground truth data for the model's training and evaluation. The team's next step in the research process was to collect a collection of images that focused on the activities of older people. These sets of images were then carefully examined. These investigations sought to elucidate significant patterns and insights concealed in the images, enabling a thorough comprehension of the behavioral characteristics associated with various levels of activity in the elderly population. Using methods designed for elderly activities, the gathered image collection was carefully organized. The images had to be resized to a constant resolution, the pixel values had to be normalized, and image enhancement techniques had to be used to boost the quality of the images. Additionally, relevant visual features that capture important data pertaining to the particular nuances of elderly activities were extracted using feature extraction techniques that were especially created for image data. These features that were extracted served as the basis for further analysis and modelling.

The ResNet50 architecture, a potent CNN model known for its efficiency in image classification tasks, is used to create the machine learning model for activity level prediction. To identify common image features and patterns, the ResNet50 model pre-trained on a sizable image dataset (like ImageNet). The final layers of ResNet50 changed to match the classification of activity levels as the desired output to adapt ResNet50 to the activity level prediction task. In order to accurately predict the corresponding activities, the model gains the ability to extract pertinent features and patterns from the images during training. The validation set is used to assess the model's performance, and if further corrections are required, such as regularization or adjusting the learning rate, they are made. Finally, the generalization ability and overall predictive accuracy of the trained ResNet50 model for activity level prediction are evaluated on the independent testing set. The assessment's findings showed how effective the methodology was at improving the ALP model's ability to identify particular activity patterns in elderly people.

The model's improvements that were seen after being trained using data on elderly people's activities highlighted the

importance of domain-specific training and adaptation to achieve the best results in activity recognition tasks for this particular population. Insightful information about the effectiveness of the methodology in identifying activities for elderly people was provided during the study's assessment phase. The study showed the effectiveness of the approach in increasing the accuracy of activity recognition for the elderly population by contrasting the model's performance before and after training and evaluating its generalization capabilities. Through the use of activity recognition technology, these findings support the creation of more dependable and efficient systems for monitoring and promoting the well-being of elderly people.

IV. RESULTS AND DISCUSSION

A) Elder Alone Detection and Device Activation

The purpose of the EAD model developed in this study was to accurately identify situations where an elder person is alone using image classification techniques. Utilizing MobileNet, effective and efficient neural network architecture, the study aimed at promoting elder safety, aiding caregivers and family members in monitoring their loved ones. The EAD model exhibited high proficiency in achieving this task, as indicated by various performance metrics. For instance, the accuracy metric demonstrated that the model correctly predicted the majority of the image classifications. This implies that the model was very successful in identifying whether an elder person was alone or not, which directly serves its fundamental purpose. The application of the MobileNet model, the preprocessing and augmentation of the dataset, and the fine-tuning of hyperparameters were key contributors to the model's effectiveness.

Due to the methodology adopted for the EAD component, successful outcomes have been achieved. It was instrumental in establishing an efficient model that not only fulfills its intended purpose but also leading to potential applications in elder care and monitoring. The implications of this study are promising advancements in the application of AI for elder safety. The Figure 2 shows the model loss after each training steps which was 6 epochs.

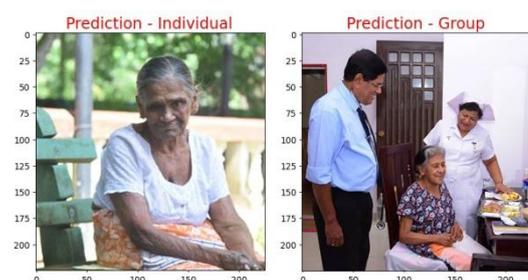


Figure 2: EAD model results output - Individual, Group

B) Automatic Accident Detection of Elderly Person

The AAD model was built to recognize the fall accident of a home alone elderly person. Initially, this model was trained for an adult dataset and it achieved 92.76% accuracy. Since this model performed well in these implementations and capability to develop in a mobile application alerting system and this was conducted with the Inception V3 model. The AAD model also trains for the Sri Lankan elderly people fall and non-fall activity datasets which are collected from local houses, and it increases the prediction response rate of the model in the Sri Lankan elderly people. Figure 3 shows the test results output of the model by test dataset.



Figure 3: AAD model results output - Fall, Not Fall

The AAD model was trained by inputting image datasets, and Figure 4 shows prediction accuracy rate (92.76%) that the model achieved. The model effective training approach and accurate test results contribute to its success and it can use for future advancements in elderly accident detection.

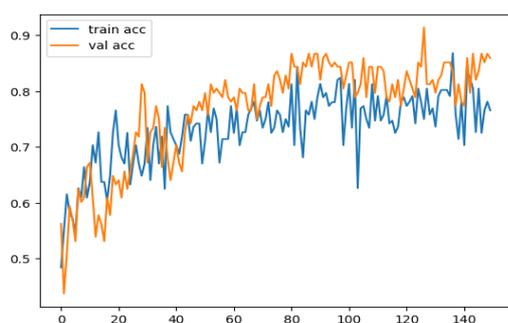


Figure 4: AAD model performance with training history graph –train acc, val acc

C) Speech Recognition with Enhanced Elderly Voice

The purpose of the EEVR model was to efficiently understand and interpret the spoken language of elderly individuals. Throughout each step of the methodology, the intent was to make the system to be more responsive and accurate towards elderly voices. One of the critical metrics to assess the performance of a speech recognition system is the Word Error Rate (WER), which represents the ratio of incorrect words recognized by the system compared to the

total number of words spoken. After the fine-tuning process with the elderly voice dataset, a substantial improvement in the model's performance was observed. The WER reduced to less than 18%, indicating that EEVR model recognized and understood the majority of the words accurately.

This decrease in the WER can be attributed to the approach, which involved feature extraction using Librosa, transfer learning, and intent classification using SVM. By training the model with an elderly voice dataset, the system became more adept at recognizing the unique characteristics present in the speech of elderly individuals, thereby improving its accuracy. In conclusion, this approach to fine-tuning the ASR model for elderly voices yielded a significant improvement in its performance.

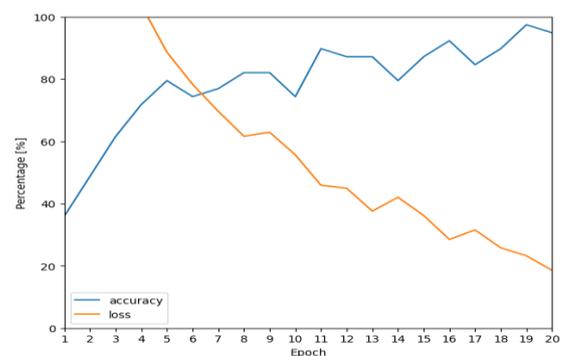


Figure 5: Accuracy and loss of the EEVR model (percentage)

D) Activity Level Prediction of Elder

The ALP model was initially trained on a dataset with the goal of identifying activities and forecasting activity levels in elderly people living alone at home. The model's proficiency in identifying the various activities carried out by seniors emphasizes how well it recognizes activities for this particular population. The ResNet50 CNN model was used in the research's activity level prediction component to predict the activity levels of elderly people. The ALP model predicted activity levels based on behavioral patterns with a significant accuracy of 91.7%. This accomplishment demonstrates the approach's effectiveness in accurately assessing and monitoring the activity levels of elderly people.



Figure 6: ALP model results output –Sitting and Standing

The behavioral pattern analysis-based ALP model was trained not only on a diverse dataset obtained from Kaggle, but also on a dataset specific to Sri Lankan elderly individuals. This strategy made sure the model was exposed to a wide variety of activity patterns from various sources, including the more general Kaggle dataset and the particular dataset from Sri Lankan elderly people. The model's performance and generalizability in predicting activity levels for elderly people were improved by including these datasets in the training process, considering both the cultural and environmental contexts of elderly people and the larger population represented in the Kaggle dataset. The collection of all activities, including standing, walking, sitting, sleeping and object-based activities from local households improved the prediction response rate of the model designed specifically for elderly Sri Lankans. The output of the model's test using the designated test dataset is shown in Figure 6. In terms of outcomes, this method was extremely successful. The ALP model was trained using image datasets, and Figure 7 shows the model's prediction accuracy rate of 91.7%.

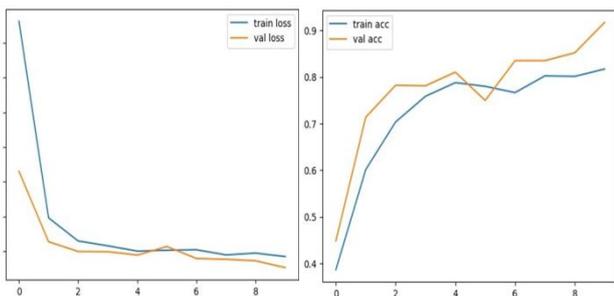


Figure 7: ALP model performance graph

In combination, the integration of these derived models culminates in a powerful and comprehensive monitoring system that boasts an outstanding overall accuracy level. This amalgamation of advanced technologies contributes to creating a highly effective and efficient monitoring device and mobile application for caregivers to personalize and closely monitor the elderly. With an overall accuracy level that consistently exceeds 90%, the Wise Eye Elderly Monitoring System ensures the safety, well-being, and overall quality of life for senior citizens, while providing caregivers with real-time insights and timely alerts, ultimately fostering a more proactive and responsive approach to elderly care.

Following the rigorous stages of model training and fine-tuning, the ensuing phase involves subjecting these refined models to comprehensive testing through hosted custom-made APIs. This dynamic process facilitates the acquisition of real-time outcomes, rendering accurate predictions across the spectrum of model functionalities. Fig 8 shows the model predictions through the hosted APIs.

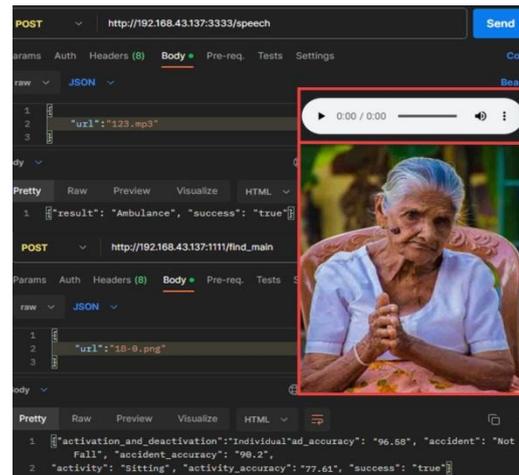


Figure 8: Four models' prediction output via hosted API testing

V. CONCLUSION

This research study unveiled the Wise Eye Elderly Monitoring System, a cutting-edge technology that uses machine learning to improve the observation and care of senior citizens. The system includes a number of crucial elements, such as automatic activation and deactivation by image classification, automatic accident detection with location identification, voice and sound recognition, and activity level prediction using behavioral pattern analysis. The Mobile net model was able to achieve 92.44% accuracy in whether an elder person was alone. Using the collected video recordings, the Inceptionv3 model was implemented to predict accident of elder: fall or not fall with an accuracy of 92.76%. Furthermore, the Speech Recognition model was possible to limit the word error rate to less than 18%. Moreover, the Resnet50 model predicts the activity levels of elderly individuals by achieving an accuracy rate of 88.44% in predicting different actions. These derived models will be used in an integrated monitoring device and mobile application for the caregivers to monitor the elderly personal.

Furthermore, privacy measures and data security protocols will be implemented to instill trust among users. Creating an open and collaborative ecosystem for researchers and healthcare professionals will foster innovation and accelerate advancements in elderly care technology. Lastly the Wise Eye elderly monitoring system will aim to pioneer the use of machine learning for long-term health prediction in elderly individuals, facilitating proactive healthcare management and improved quality of life. Overall, this system is benefiting countless seniors and their caregivers in the journey towards a safer and healthier aging experience.

REFERENCES

- [1] Boukhenoufa, A. Amira, F. Bensaali, and S. Soheilian Esfahani, "A novel gateway-based solution for remote

- elderly monitoring," Journal of Biomedical Informatics, vol. 109, Sep. 2020, article no. 103521, doi: 10.1016/j.jbi.2020.103521. PMID: 32745621.
- [2] Y. Karunanayake, U. Thayasivam, and S. Ranathunga, "Transfer Learning Based Free-Form Speech Command Classification for Low-Resource Languages," in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, Florence, Italy, 2019, pp. 288-294.
- [3] Y. Karunanayake, U. Thayasivam and S. Ranathunga, "Sinhala and Tamil Speech Intent Identification From English Phoneme Based ASR," 2019 International Conference on Asian Language Processing (IALP), Shanghai, China, 2019, pp. 234-239, doi: 10.1109/IALP48816.2019.9037702.
- [4] F. Delmastro, F. Di Martino, and C. Dolciotti, "Cognitive Training and Stress Detection in MCI Frail Older People Through Wearable Sensors and Machine Learning," IEEE Access, vol. 8, pp. 65573-65590, 2020.
- [5] T. S. Barger, D. E. Brown, and M. Alwan, "Health-status monitoring through analysis of behavioral patterns," IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 35, no. 1, pp. 22-27, Jan. 2005, doi: 10.1109/TSMCA.2004.838474.
- [6] S. Deepika and K. P. Vijayakumar, "IoT based Elderly Monitoring System," presented at the 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 573-579. doi: 10.1109/ICOEI53556.2022.9777186.
- [7] S. Anderson et al., "Recognition of elderly speech and voice-driven document retrieval," in 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258), Phoenix, AZ, USA, 1999, pp. 145-148 vol.1, doi: 10.1109/ICASSP.1999.758083.
- [8] V. W. T. N. a. R. W. H. Karunathilaka, "Low-resource Sinhala Speech Recognition using Deep Learning," in International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, 2020.
- [9] M. Fukuda, H. Nishizaki, Y. Iribe, R. Nishimura and N. Kitaoka, "Improving Speech Recognition for the Elderly: A New Corpus of Elderly Japanese Speech and Investigation of Acoustic Modeling for Speech Recognition," in 12th Conference on Language Resources and Evaluation (LREC 2020), Marseille, 2020.
- [10] "Human Action (7 Actions)" [Online] Available : <https://www.kaggle.com/datasets/arghudas2021/human-action-7-actions>
- [11] "Human Action Recognition (HAR) Dataset" [Online] Available : <https://www.kaggle.com/datasets/meetnagadia/human-action-recognition-har-dataset>
- [12] "Keras," [Online]. Available: <https://keras.io/api/applications/mobilenet/>. [Accessed 07 2023].
- [13] "Fall Detection Dataset" [Online] Available: <https://www.kaggle.com/datasets/uttejkumarkandagatla/fall-detection-dataset>
- [14] "Fall-detection-classification" [Online] Available: <https://www.kaggle.com/datasets/vishnu6174/falldetectionclassification>
- [15] "Clean-Elderly-Dataset-v1-fall-detection" [Online] Available: <https://www.kaggle.com/datasets/sreevishnukumarv1/cleanelderlydatasetv1falldetection>
- [16] "Fall Detection dataset" [Online] Available: <https://www.kaggle.com/datasets/bernardngangbam/fall-detection>
- [17] "Dataset CAUCAFALL" [Online] Available: <https://data.mendeley.com/datasets/7w7fccy7ky/4>
- [18] "KFall: A Comprehensive Motion Dataset to Detect Pre-impact Fall for the Elderly based on Wearable Inertial Sensors" [Online] Available : <https://sites.google.com/view/kfalldataset?pli=1>

Citation of this Article:

Amila Sampath, Akila Kavinda, Sachintha Gunaratne, Thushal Shaminda, "WiseEye Elderly Monitoring System Using Machine Learning" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 11, pp 35-41, November 2023. Article DOI <https://doi.org/10.47001/IRJIET/2023.711006>
